#### SEQUENCE LISTING

#### (1) GENERAL INFORMATION:

- (i) APPLICANT: POTTER, ANDREW A. REDMOND, MARK J. HUGHES, HUW P.A.
- (ii) TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN CHIMERAS
- (iii) NUMBER OF SEQUENCES: 15
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: REED & ROBINS
  - (B) STREET: 285 HAMILTON AVENUE, SUITE 200
  - (C) CITY: PALO ALTO
  - (D) STATE: CALIFORNIA
  - (E) COUNTRY: UNITED STATES OF AMERICA
  - (F) ZIP: 94301
  - (v) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: Floppy disk
    - (B) COMPUTER: IBM PC compatible
    - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/455,970
  - (B) FILING DATE: 31-MAY-1995
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/960,932
  - (B) FILING DATE: 14-OCT-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: ROBINS, ROBERTA L.
  - (B) REGISTRATION NUMBER: 33,208
  - (C) REFERENCE/DOCKET NUMBER: 9001-0016.10
  - (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: (415) 327-3400
    - (B) TELEFAX: (415) 327-3231
- (2) INFORMATION FOR SEQ ID NO:1:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 2794 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (ix) FEATURE:
    - (A) NAME/KEY: CDS
    - (B) LOCATION: 1..2778

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG Met 1	GCT Ala	ACT Thr	GTT Val	ATA Ile 5	GAT Asp	CTA Leu	AGC Ser	TTC Phe	CCA Pro 10	AAA Lys	ACT Thr	GGG Gly	GCA Ala	AAA Lys 15	AAA Lys	48
ATT Ile	ATC Ile	CTC Leu	TAT Tyr 20	ATT Ile	CCC Pro	CAA Gln	AAT Asn	TAC Tyr 25	CAA Gln	TAT Tyr	GAT Asp	ACT Thr	GAA Glu 30	CAA Gln	GGT Gly	96
AAT Asn	GGT Gly	TTA Leu 35	CAG Gln	GAT Asp	TTA Leu	GTC Val	AAA Lys 40	GCG Ala	GCC Ala	GAA Glu	GAG Glu	TTG Leu 45	GGG Gly	ATT Ile	GAG Glu	144
GTA Val	CAA Gln 50	AGA Arg	GAA Glu	GAA Glu	CGC Arg	AAT Asn 55	AAT Asn	ATT Ile	GCA Ala	ACA Thr	GCT Ala 60	CAA Gln	ACC Thr	AGT Ser	TTA Leu	192
GGC Gly 65	Thr	ATT Ile	CAA Gln	ACC Thr	GCT Ala 70	ATT Ile	GGC Gly	TTA Leu	ACT Thr	GAG Glu 75	CGT Arg	GGC Gly	ATT Ile	GTG Val	TTA Leu 80	240
TCC Ser	GCT Ala	CCA Pro	CAA Gln	ATT Ile 85	Asp	AAA Lys	TTG Leu	CTA Leu	CAG Gln 90	гAг	ACT Thr	AAA Lys	GCA Ala	GGC Gly 95	CAA Gln	288
GCA Ala	TTA Leu	GGT	TCT Ser 100	Ala	GAA Glu	AGC Ser	ATT	GTA Val 105	GIII	AAT Asn	GCA Ala	AAT Asn	AAA Lys 110	GCC Ala	AAA Lys	336
ACT Thr	GTA Val	TT <i>I</i> Lev	ı Sei	GGC Gly	ATT	CAA Gln	TCT Ser 120	tre	TTA Leu	GGC Gly	TCA Ser	GTA Val 125	neu	GCT Ala	GGA Gly	384
ATC Met	GAT Asg 130	Le	A GA:	GAC Glu	GCC Ala	TTA Lev 135	GII	AAT Asn	AAC Asn	AGC Ser	AAC Asr 140	GII	CAT His	GCT Ala	CTT Leu	432
GC: Ala 14!	a Lys	A GC'	r GGG a Gl	TTO Y Lev	G GAC 1 Glu 150	ı Let	ACI 1 Thi	AA A C Asr	TCA n Ser	155	1 110	GAA Glu	AAT Asn	ATT	GCT Ala 160	480
AA' As:	r TC	A GT. r Va	A AA l Ly	A ACI	r Lei	r GA( 1 Asj	GAI Glu	A TTT	r GGT e Gly 170	7 610	G CAI	A ATT	AG1 Ser	CAA Glr 175	TTT Phe	528
GG Gl	т тс y Se	A AA r Ly	A CT s Le 18	u GI	A AA' n As:	r ATO	C AA e Ly	A GGG s Gl	y ne	A GGG	G AC' y Th	r TTI	A GG/ 1 Gly 190		Lys	576
CT Le	C AA u Ly	A AA s As	n Il	C GG e Gl	т GG y Gl	A CT y Le	T GA u As 20	Бгй	A GC	T GG a Gl	C CT y Le	T GG' u Gl; 20	,	A GA' u Asj	r GTT o Val	624
AT Il	C TC e Se 21	r Gl	G CI	A TI eu Le	'A TC	G GG r Gl 21	У АТ	A AC a Th	A GC r Al	T GC a Al	A CT a Le 22	<u>u</u> •u	A CT	T GC	A GAT a Asp	672
A# Ly 22	s As	T GO	CT TO La Se	CA AC	CA GC nr Al 23	а ьу	A AA s Ly	A GT 's Va	rg gg il gl	T GC y Al 23	.a 01	T TT. Y Ph	T GA e Gl	A TT u Le	G GCA u Ala 240	

· · · · · · · · · · · · · · · · · · ·				•					V.	1 .,	•						
AAC Asn	CAA Gln	GTT Val	GTT Val	GGT Gly 245	AAT Asn	ATT :	ACC :	Lys	GCC Ala 250	GTT Val	TCT Ser	TCT Ser	TAC Tyr	ATT Ile 255	TTA Leu	768	
GCC Ala	CAA Gln	CGT Arg	GTT Val 260	GCA Ala	GCA Ala	GGT Gly	Leu	TCT Ser 265	TCA Ser	ACT Thr	GGG Gly	CCT Pro	GTG Val 270	GCT Ala	GCT Ala	816	
TTA Leu	ATT Ile	GCT Ala 275	TCT Ser	ACT Thr	GTT Val	TCT Ser	CTT Leu 280	GCG Ala	ATT Ile	AGC Ser	CCA Pro	TTA Leu 285	GCA Ala	TTT Phe	GCC Ala	864	
GGT Gly	ATT Ile 290	Ala	GAT Asp	AAA Lys	TTT Phe	AAT Asn 295	CAT His	GCA Ala	AAA Lys	AGT Ser	TTA Leu 300	GAG Glu	AGT Ser	TAT Tyr	GCC Ala	912	
GAA Glu 305	CGC Arg	TTT Phe	AAA Lys	AAA Lys	TTA Leu 310	GGC Gly	TAT Tyr	GAC Asp	GGA Gly	GAT Asp 315	AAT Asn	TTA Leu	TTA Leu	GCA Ala	GAA Glu 320	960	
TAT Tyr	CAG	CGG Arg	GGA Gly	ACA Thr 325	GGG Gly	ACT Thr	ATT Ile	GAT Asp	GCA Ala 330	TCG Ser	GTT Val	ACT Thr	GCA Ala	ATT Ile 335	AAT Asn	1008	
ACC Thr	GCA Ala	TTG Leu	GCC Ala 340	Ala	ATT Ile	GCT Ala	GGT Gly	GGT Gly 345	GTG Val	TCT Ser	GCT Ala	GCT Ala	GCA Ala 350	GCC Ala	GGC Gly	1056	
Ser	Va]	1 11e 355	Ala	Ser	Pro		360	Leu	Leu	vai	261	365			017	1104	
Val	370	e Ser	Thr	: Ile	Leu	375	Tyr	ser	гур	GIII	380	Nec	1	014		1152	
Val 385	Al	a Ası	ı Lys	; Ile	390	Asn	гуѕ	IIe	· Vai	395		010	יי בי		AAT Asn 400	1200	
His	s Gl	у Гу	s Ası	1 Tyr	Phe	e Glu	ASD	. Сту	410	)	, AIC	, mag	, -,-	415		1248	
Ası	ı Le	u Gl:	n Asj 42	p Ası O	n Met	: Lys	Pne	425	i Let	ı ASı	, nec	. ASI	430		TTA Leu	1296	
Gli	n Al	a Gl 43	u Ar	g Va.	1 110	e Ala	440	)	C GII		. 011	44	5		n Asn	1344	
Il	e G] 45	y As 50	p Le	u Al	a GI	y 116 459	s Sei	. Ari	а пе	u Gi	46	0			r AGT u Ser	1392	
Gl	T A/ y L) 5	AA GC /s Al	C TA a Ty	T GT T Va	G GA 1 As 47	b AT	TT:	r GA e Gl	A GA u Gl	A GG u Gl 47	y y	A CA s Hi	C AT s Il	T AA e Ly	A GCC s Ala 480	1440	

GAT Asp	AAA Lys	TTA Leu	GTA Val	CAG Gln 485	TTG Leu	GAT Asp	TCG Ser	GCA Ala	AAC Asn 490	GGT Gly	ATT Ile	ATT Ile	GAT Asp	GTG Val 495	AGT Ser	1488
AAT Asn	TCG Ser	GGT Gly	AAA Lys 500	GCG Ala	AAA Lys	ACT Thr	CAG Gln	CAT His 505	ATC Ile	TTA Leu	TTC Phe	AGA Arg	ACG Thr 510	CCA Pro	TTA Leu	1536
TTG Leu	ACG Thr	CCG Pro 515	GGA Gly	ACA Thr	GAG Glu	CAT His	CGT Arg 520	GAA Glu	CGC Arg	GTA Val	CAA Gln	ACA Thr 525	GGT Gly	AAA Lys	TAT Tyr	1584
GAA Glu	TAT Tyr 530	ATT Ile	ACC Thr	AAG Lys	CTC Leu	AAT Asn 535	ATT Ile	AAC Asn	CGT Arg	GTA Val	GAT Asp 540	AGC Ser	TGG Trp	AAA Lys	ATT Ile	1632
ACA Thr 545	GAT Asp	GGT Gly	GCA Ala	GCA Ala	AGT Ser 550	TCT Ser	ACC Thr	TTT Phe	GAT Asp	TTA Leu 555	ACT Thr	AAC Asn	GTT Val	GTT Val	CAG Gln 560	1680
CGT Arg	ATT Ile	GGT Gly	ATT	GAA Glu 565	TTA Leu	GAC Asp	AAT Asn	GCT Ala	GGA Gly 570	AAT Asn	GTA Val	ACT Thr	AAA Lys	ACC Thr 575	AAA Lys	1728
GAA Glu	ACA Thr	AAA Lys	ATT	Ile	GCC Ala	AAA Lys	CTT Leu	GGT Gly 585	GIU	GGT Gly	GAT Asp	GAC Asp	AAC Asn 590	GTA Val	TTT Phe	1776
GTT Val	GGT Gly	TCT Ser	: Gly	ACG Thr	ACG Thr	GAA Glu	ATT Ile 600	Asp	GGC Gly	GGT Gly	GAA Glu	GGT Gly 605	- 1 -	GAC Asp	CGA Arg	1824
GTT Val	CAC His	Ty	r AGG	C CGT	GGA Gly	AAC Asn 615	Tyr	GGT Gly	GCT Ala	TTA Leu	ACT Thr 620		GAT Asp	GCA Ala	ACC Thr	1872
AAI Lys 625	s Glu	AC 1 Th	C GAG	G CAZ u Glr	GGT Gly 630	sei	TAT Tyr	ACC Thi	GT#	A AAT Asn 635		TTC J Phe	GTA Val	GAA Glu	ACC Thr 640	1920
Gl	у Гу	s Al	a Le	u Hi:	s GIV 5	ı val	LTM	. se	650	)	, 111			655		1968
As	n Ar	g Gl	u G1 66	u Ly O	S 116	S GT	ı ıy.	66	5	<i>y</i> DC.			670	0	C CAT s His	2016
Al	a Gl	у Ту 67	r Ty 5	r Th	r Ly	s As	68 68	O Te	и шу	S AL	u vu	68	5		T ATC e Ile	
Gl	y Th	r Se	er Hi	s As	n As	p 11 69	e Pn 5	е гу	5 61	у эс	70	0			T GCC p Ala	
TI Ph 70	ne As	C G(sn G)	er Go ly G	GT GA Ly As	T GG sp Gl 71	y va	C GA 1 As	T AC	T AT	T GA e As 71	5 07	T AA .y As	C GA	C GG p Gl	C AAT y Asn 720	2160

GAC Asp	CGC Arg	TTA Leu	Phe	GGT Gly 725	GGT Gly	AAA Lys	GGC Gly	GAT Asp	GAT Asp 730	ATT Ile	CTC Leu	GAT Asp	GGT Gly	GGA Gly 735	AAT Asn	2208
GGT Gly	GAT Asp	GAT Asp	TTT Phe 740	ATC Ile	GAT Asp	GGC Gly	GGT Gly	AAA Lys 745	GGC Gly	AAC Asn	GAC Asp	Leu	TTA Leu 750	CAC His	GGT Gly	2256
GGC Gly	AAG Lys	GGC Gly 755	GAT Asp	GAT Asp	ATT Ile	TTC Phe	GTT Val 760	CAC His	CGT Arg	AAA Lys	GGC Gly	GAT Asp 765	GGT Gly	AAT Asn	GAT Asp	2304
ATT Ile	ATT Ile 770	ACC Thr	GAT Asp	TCT Ser	GAC Asp	GGC Gly 775	AAT Asn	GAT Asp	AAA Lys	TTA Leu	TCA Ser 780	TTC Phe	TCT Ser	GAT Asp	TCG Ser	2352
AAC Asn 785	TTA Leu	AAA Lys	GAT Asp	TTA Leu	ACA Thr 790	TTT Phe	GAA Glu	AAA Lys	GTT Val	AAA Lys 795	CAT	AAT Asn	CTT Leu	GTC Val	ATC Ile 800	2400
ACG Thr	AAT Asn	AGC Ser	AAA Lys	AAA Lys 805	GAG Glu	AAA Lys	GTG Val	ACC Thr	ATT Ile 810	CAA Gln	AAC Asn	TGG Trp	TTC Phe	CGA Arg 815	GAG Glu	2448
GCT Ala	GAT Asp	TTT Phe	GCT Ala 820	AAA Lys	GAA Glu	GTG Val	CCT Pro	AAT Asn 825	TAT Tyr	AAA Lys	GCA Ala	ACT Thr	AAA Lys 830	GAT Asp	GAG Glu	2496
AAA Lys	ATC Ile	GAA Glu 835	Glu	ATC Ile	ATC Ile	GGT Gly	CAA Gln 840	Asn	GGC Gly	GAG Glu	CGG Arg	ATC Ile 845	ACC Thr	TCA Ser	AAG Lys	2544
CAA Gln	GTT Val 850	Asp	GAT Asp	CTT Leu	ATC Ile	GCA Ala 855	. Lys	GGT Gly	AAC Asn	GGC Gly	AAA Lys 860	116	ACC Thr	CAA Gln	GAT Asp	2592
GAG Glu 865	ı Lev	TCA Ser	AAA : Lys	GTT Val	GTT Val 870	Asp	AAC Asr	TAT Tyr	GAA Glu	TTG Leu 875	, nec	AAA Lys	CAT His	AGC Ser	Lys 880	2640
Ası	ı Val	LThi	. Asr	ser 885	Lev	ı Asr	) ràs	s Let	890	)	. 561	. vai		895		2688
ACC Th	TCC r Sea	TC:	T AAT ASI 900	ı Ası	TCC Sei	AGA	A AA:	GT/ n Val 90!	ь пе	A GTO	GCT L Ala	r CCA a Pro	ACT Thr 910		A ATG	2736
TT( Le	G GA' u Asj	r CA p Gl: 91	n Se	r TTI	A TC:	r TC	r CT' r Le	r CA u Gli 0	A TT'	r GC e Ala	r AGG	G GG/ g Gly 929	261	C r		2778
TA	GCTA	GCTA	GCC	ATG												2794

### (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 926 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
- Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys

  10 15
- Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly
  20 25 30
- Asn Gly Leu Gln Asp Leu Val Lys Ala Ala Glu Glu Leu Gly Ile Glu 35 40 45
- Val Gln Arg Glu Glu Arg Asn Asn Ile Ala Thr Ala Gln Thr Ser Leu
  50 55 60
- Gly Thr Ile Gln Thr Ala Ile Gly Leu Thr Glu Arg Gly Ile Val Leu
  65 70 75 80
- Ser Ala Pro Gln Ile Asp Lys Leu Leu Gln Lys Thr Lys Ala Gly Gln 85 90 95
- Ala Leu Gly Ser Ala Glu Ser Ile Val Gln Asn Ala Asn Lys Ala Lys
- Thr Val Leu Ser Gly Ile Gln Ser Ile Leu Gly Ser Val Leu Ala Gly 115 120 125
- Met Asp Leu Asp Glu Ala Leu Gln Asn Asn Ser Asn Gln His Ala Leu
- Ala Lys Ala Gly Leu Glu Leu Thr Asn Ser Leu Ile Glu Asn Ile Ala 145 150 150 160
- Asn Ser Val Lys Thr Leu Asp Glu Phe Gly Glu Gln Ile Ser Gln Phe 165 170 175
- Gly Ser Lys Leu Gln Asn Ile Lys Gly Leu Gly Thr Leu Gly Asp Lys 180 185 190
- Leu Lys Asn Ile Gly Gly Leu Asp Lys Ala Gly Leu Gly Leu Asp Val 195 200 205
- Ile Ser Gly Leu Leu Ser Gly Ala Thr Ala Ala Leu Val Leu Ala Asp 210 220
- Lys Asn Ala Ser Thr Ala Lys Lys Val Gly Ala Gly Phe Glu Leu Ala 225 230 235 240
- Asn Gln Val Val Gly Asn Ile Thr Lys Ala Val Ser Ser Tyr Ile Leu 245 250 255
- Ala Gln Arg Val Ala Ala Gly Leu Ser Ser Thr Gly Pro Val Ala Ala 260 265 270

Leu Ile Ala Ser Thr Val Ser Leu Ala Ile Ser Pro Leu Ala Phe Ala 275 280 285

Gly Ile Ala Asp Lys Phe Asn His Ala Lys Ser Leu Glu Ser Tyr Ala 290 295 300

Glu Arg Phe Lys Lys Leu Gly Tyr Asp Gly Asp Asn Leu Leu Ala Glu 305 310 315 320

Tyr Gln Arg Gly Thr Gly Thr Ile Asp Ala Ser Val Thr Ala Ile Asn 325 330 335

Thr Ala Leu Ala Ala Ile Ala Gly Gly Val Ser Ala Ala Ala Gly 340 345 350

Ser Val Ile Ala Ser Pro Ile Ala Leu Leu Val Ser Gly Ile Thr Gly 355 360 365

Val Ile Ser Thr Ile Leu Gln Tyr Ser Lys Gln Ala Met Phe Glu His 370 380

Val Ala Asn Lys Ile His Asn Lys Ile Val Glu Trp Glu Lys Asn Asn 385 390 395 400

His Gly Lys Asn Tyr Phe Glu Asn Gly Tyr Asp Ala Arg Tyr Leu Ala 405 410 415

Asn Leu Gln Asp Asn Met Lys Phe Leu Leu Asn Leu Asn Lys Glu Leu 420 425 430

Gln Ala Glu Arg Val Ile Ala Ile Thr Gln Gln Gln Trp Asp Asn Asn 435

Ile Gly Asp Leu Ala Gly Ile Ser Arg Leu Gly Glu Lys Val Leu Ser

Gly Lys Ala Tyr Val Asp Ala Phe Glu Glu Gly Lys His Ile Lys Ala 465 470 475 480

Asp Lys Leu Val Gln Leu Asp Ser Ala Asn Gly Ile Ile Asp Val Ser

Asn Ser Gly Lys Ala Lys Thr Gln His Ile Leu Phe Arg Thr Pro Leu
500 505 510

Leu Thr Pro Gly Thr Glu His Arg Glu Arg Val Gln Thr Gly Lys Tyr 515 520 525

Glu Tyr Ile Thr Lys Leu Asn Ile Asn Arg Val Asp Ser Trp Lys Ile

Thr Asp Gly Ala Ala Ser Ser Thr Phe Asp Leu Thr Asn Val Val Gln 545 550 560

Arg Ile Gly Ile Glu Leu Asp Asn Ala Gly Asn Val Thr Lys Thr Lys 565 570 575

Glu Thr Lys Ile Ile Ala Lys Leu Gly Glu Gly Asp Asp Asn Val Phe

Val Gly Ser Gly Thr Thr Glu Ile Asp Gly Gly Glu Gly Tyr Asp Arg 595 600 605

Val His Tyr Ser Arg Gly Asn Tyr Gly Ala Leu Thr Ile Asp Ala Thr Lys Glu Thr Glu Gln Gly Ser Tyr Thr Val Asn Arg Phe Val Glu Thr Gly Lys Ala Leu His Glu Val Thr Ser Thr His Thr Ala Leu Val Gly 650 Asn Arg Glu Glu Lys Ile Glu Tyr Arg His Ser Asn Asn Gln His His 665 Ala Gly Tyr Tyr Thr Lys Asp Thr Leu Lys Ala Val Glu Glu Ile Ile Gly Thr Ser His Asn Asp Ile Phe Lys Gly Ser Lys Phe Asn Asp Ala 695 Phe Asn Gly Gly Asp Gly Val Asp Thr Ile Asp Gly Asn Asp Gly Asn Asp Arg Leu Phe Gly Gly Lys Gly Asp Asp Ile Leu Asp Gly Gly Asn Gly Asp Asp Phe Ile Asp Gly Gly Lys Gly Asn Asp Leu Leu His Gly Gly Lys Gly Asp Asp Ile Phe Val His Arg Lys Gly Asp Gly Asn Asp Ile Ile Thr Asp Ser Asp Gly Asn Asp Lys Leu Ser Phe Ser Asp Ser Asn Leu Lys Asp Leu Thr Phe Glu Lys Val Lys His Asn Leu Val Ile Thr Asn Ser Lys Lys Glu Lys Val Thr Ile Gln Asn Trp Phe Arg Glu Ala Asp Phe Ala Lys Glu Val Pro Asn Tyr Lys Ala Thr Lys Asp Glu 825 -Lys Ile Glu Glu Ile Ile Gly Gln Asn Gly Glu Arg Ile Thr Ser Lys Gln Val Asp Asp Leu Ile Ala Lys Gly Asn Gly Lys Ile Thr Gln Asp 855 Glu Leu Ser Lys Val Val Asp Asn Tyr Glu Leu Leu Lys His Ser Lys Asn Val Thr Asn Ser Leu Asp Lys Leu Ile Ser Ser Val Ser Ala Phe Thr Ser Ser Asn Asp Ser Arg Asn Val Leu Val Ala Pro Thr Ser Met 905

Leu Asp Gln Ser Leu Ser Ser Leu Gln Phe Ala Arg Gly Ser

(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 60 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
GATCCAGCTC TTCTGCCGGC TGCAAAAACT TCTTCTGGAA AACCTTCACC AGCTGCTAGG	60
(2) INFORMATION FOR SEQ ID NO:4:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 60 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
GATCCCTAGC AGCTGGTGAA GGTTTTCCAG AAGAAGTTTT TGCAGCCGGC AGAAGAGCTG	60
(2) INFORMATION FOR SEQ ID NO:5:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 39 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
GATCTCAGCA TTGGAGCTAC GGCCTGCGCC CTGGCTAAG	39
(2) INFORMATION FOR SEQ ID NO:6:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 39 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
GATCCTTAGC CAGGGCGCAG GCCGTAGCTC CAATGCTGA	3

(2) INFORMATION FOR SEQ ID NO:3:

(2) INFORMATION FOR SEQ ID NO:7:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 83 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GATCTTGCAA CATTGTGCCT GTGAGCATTG TGAGCCGCAA CATTGTGTAC ACCCGCGCGC	60
AACCTAACCA AGACATTGTG TAG	83
(2) INFORMATION FOR SEQ ID NO:8:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 83 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
GATCCTACAC AATGTCTTGG TTAAGTTGCG CGCGGGTGTA CACAATGTTG CGGCTCACAA	60
TCGTCACAGG CACAATGTTG CAA	83
(2) INFORMATION FOR SEQ ID NO:9:	
<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 2838 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: double</li> <li>(D) TOPOLOGY: linear</li> </ul>	
(ii) MOLECULE TYPE: DNA (genomic)	
(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 12829	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	
ATG GCT ACT GTT ATA GAT CTA AGC TTC CCA AAA ACT GGG GCA AAA AAA Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys 1 5 10 15	48
ATT ATC CTC TAT ATT CCC CAA AAT TAC CAA TAT GAT ACT GAA CAA GGT Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly 20 25 30	96
AAT GGT TTA CAG GAT TTA GTC AAA GCG GCC GAA GAG TTG GGG ATT GAG Asn Gly Leu Gln Asp Leu Val Lys Ala Ala Glu Glu Leu Gly Ile Glu 35 40 45	144

																		•
				•		•				ζ' - '								
GTA Val	CAA Gli	n A	AGA Arg	GAA Glu	GAA Glu	CGC Arg	AAT Asn 55	AAT . Asn	ATT	GCA Ala	ACA Thr	GCT Ala 60	CAA Gln	ACC Thr	AGT Ser	L.	TA eu	192
GGC Gly 65	ACC Th:	G A	ATT Ile	CAA Gln	ACC Thr	GCT Ala 70	ATT Ile	GGC Gly	TTA Leu	ACT Thr	GAG Glu 75	CGT Arg	GGC Gly	ATT	GTG Val		TA eu 80	240
TCC Ser	GC'	T (	CCA Pro	CAA Gln	ATT Ile 85	GAT Asp	AAA Lys	TTG Leu	CTA Leu	CAG Gln 90	AAA Lys	ACT Thr	AAA Lys	GCA Ala	GGC Gly 95		AA ln	288
GCA Ala	TT Le	A (	GGT Gly	TCT Ser 100	GCC Ala	GAA Glu	AGC Ser	ATT Ile	GTA Val 105	CAA Gln	AAT Asn	GCA Ala	AAT Asn	AAA Lys 110		Z A	AA .ys	336
ACT Thr	GT Va	1	TTA Leu 115	TCT Ser	GGC Gly	ATT Ile	CAA Gln	TCT Ser 120	ATT Ile	TTA Leu	GGC Gly	TCA Ser	GTA Val 125		GC' Ala	r G	GA Gly	384
ATG Met	GA As	g	TTA Leu	GAT Asp	GAG Glu	GCC Ala	TTA Leu 135	CAG Gln	AAT Asn	AAC Asn	AGC Ser	AAC Asn 140	. 011	CAT His	r GC' s Ala	r ( a I	CTT Leu	432
GCT Ala 145	AA L		GCT Ala	GGC Gly	TTO	GAG Glu 150	Leu	ACA Thr	AAT Asn	TCA Ser	TTA Leu 155		GAZ Glu	A AA' 1 Asi	r AT n Il	т ( е і	GCT Ala 160	480
		CA er	GTA Val	AAA Lys	ACA Thi	: Let	GAC Asp	GAA Glu	TTT Phe	GGT Gly 170	GIU	CAF Glr	A AT	r AG e Se	T CA r Gl 17		TTT Phe	528
GG7 Gl)	r T	CA er	AAA Lys	CTI Lev	ı Glı	A AA n Asi	r ATC	AAA Lys	GGC Gly 185	Dec	GGG Gly	ACT Thi	r TT	A GG u Gl 19	4	.C	AAA Lys	576
CT( Le	C A	AA ys	AAT Asi	ı Il	C GG e Gl	r GG y Gl	A CTT	GAT Asp 200	, nys	GCT Ala	GGG Gly	C CT	r GG u Gl 20	<b>-</b> -	'A GA u As	T	GTT Val	624
AT Il	e S	CA Ser	Gl	G CT y Le	A TT u Le	A TC u Se	G GGG r Gl	C GCA y Ala 5	A ACA	A GCT	r GC	A CT a Le 22	• • •	A CT l Le	T GO	CA La	GAT Asp	672
AA Ly 22	s P	AAT Asn	GC'	T TC a Se	A AC	A GC r Al 23	а ьу	A AA! s Ly:	A GTO	G GG' L Gl	T GC y Al 23	a 01	T TI y Ph	T GA	AA T'	rg eu	GCA Ala 240	720
		CAA Glr	A GT 1 Va	T GT 1 Va	T GG	y As	T AT	T ACC	C AA	A GC s Al 25		T TC l Se	T TO	T T	AC A' yr I 2	TT le 55	TTA Leu	768
GC Al	cc ( la (	CA <i>I</i> Glr	A CG n Ar	g Va	T GC	A GC La Al	CA GG .a Gl	T TT. y Le	A TC u Se 26	1 36	A AC	T GO	G CO Ly Pi		TG G al A 70	CT la	GCT Ala	816
TT Le	ra . eu	AT:	r GC e Al 27	a Se	er T	or Gr nr Va	TT TO	T CT er Le 28	u AI	G AT a Il	T AC	SC CO		TA G eu A 85	CA T la P	TT he	GCC Ala	864

				· •		•				· · ·	' ',	'					
GG Gl	у	ATT (	GCC Ala	GAT Asp	AAA Lys	Phe	AAT Asn 295	CAT His	GCA Ala	AAA Lys	AGT Ser	TTA Leu 300	GAG Glu	AGT Ser	TAT Tyr	GCC Ala	912
GA Gl 30	u .	CGC Arg	TTT Phe	AAA Lys	AAA Lys	TTA Leu 310	GGC Gly	TAT Tyr	GAC Asp	GGA Gly	GAT Asp 315	AAT Asn	TTA Leu	TTA Leu	GCA Ala	GAA Glu 320	960
TA Ty	T	CAG Gln	CGG Arg	GGA Gly	ACA Thr 325	GGG Gly	ACT Thr	ATT Ile	GAT Asp	GCA Ala 330	TCG Ser	GTT Val	ACT Thr	GCA Ala	ATT Ile 335	AAT Asn	1008
Tì	ır	Ala	Leu	Ala 340	Ala	Ile	Ala	GIÀ	345	vai	ser	GCT Ala	AIA	350	AIG	GIY	1056
Se	er	Val	Ile 355	Ala	Ser	Pro	Ile	360	Leu	Leu	vai	TCT Ser	365	116	1111	Gry	1104
V	al	11e 370	Ser	Thr	Ile	Leu	375	Tyr	ser	гуз	GIII	GCA Ala 380	Mec	FIIC			1152
V: 3	al 85	Ala	Asn	Lys	Ile	His 390	Asn	rys	TTE	Val	395		GIU	БуЗ	71011	400	1200
Н	is	Gly	Lys	Asn	Tyr 405	Phe	GIU	Asn	GIY	410	Asp	Ala	Arg	-7-	415	GCG Ala	1248
A	sn	Leu	Gln	420	Asn	Met	гуѕ	Pne	425	. nea	, ASI	. Dea		430		TTA Leu	1296
G	ln	Ala	Glu 435	Arg	y Val	lle	Ala	440		. GII		. 011.	445			AAC Asn	1344
3	[le	Gly 450	/ Asr	Let	ı Ala	. GIÀ	455	s ser	MIG	, nec	I GI	460	)		-	AGT Ser	1392
	31y 465	r Lys	a Ala	а Ту	r Val	470	) O	i Pne	3 610	1 61	47	5	,			A GCC S Ala 480	1440
]	Asp	Ly:	s Le	u Va	1 Gl: 48	n Let	u Asj	o se	LAI	49	0	,			49		1488
	Ası	n Se	r Gl	y Ly 50	s Ala	а ∟у	s Tn	r Gi.	50	5	C 11C	u		51	0	A TTA	1536
	TT(	G AC u Th	G CC r Pr 51	o Gl	A AC y Th	A GA r Gl	G CA u Hi	T CG s Ar 52	g Gr	A CG u Ar	C GT g Va	A CA	A AC n Th 52		T AA y Ly	A TAT s Tyr	1584

•

٠			•		•				, 4	, , , ,	•			1						
											•									
GAA ' Glu '	TAT . Tyr 530	ATT Ile	ACC Thr	AAG Lys	CTC Leu	AAT Asn 535	ATT Ile	AAC Asn	CGT ( Arg '	val A	GAT . Asp	AGC Ser	TGG Trp	A)	AA i	ATT Ile	1632			
ACA Thr	GAT Asp	GGT Gly	GCA Ala	GCA Ala	AGT Ser 550	TCT Ser	ACC Thr	TTT Phe	Asp	TTA Z Leu ' 555	ACT Thr	AAC Asn	GTT Val	' G'	aı '	CAG Gln 560	1680			
CGT Arg	ATT Ile	GGT Gly	ATT Ile	GAA Glu 565	TTA Leu	GAC Asp	AAT Asn	GCT Ala	GGA Gly 570	AAT Asn	GTA Val	ACT Thr	AAA Lys	Ι.	CC hr 75	AAA Lys ·	1728			
GAA Glu	ACA Thr	AAA Lys	ATT Ile 580	ATT Ile	GCC Ala	AAA Lys	CTT Leu	GGT Gly 585	GAA Glu	GGT Gly	GAT Asp	GAC Asp	AAC Asn 590		TA al	TTT Phe	1776			
GTT Val	GGT Gly	TCT Ser 595	GGT Gly	ACG Thr	ACG Thr	GAA Glu	ATT Ile 600	GAT Asp	GGC Gly	GGT Gly	GAA Glu	GGT Gly 605	TAC	C G	AC sp	CGA Arg	1824			
GTT Val	CAC His 610	TAT Tyr	AGC Ser	CGT Arg	GGA Gly	AAC Asn 615	ТАТ Туг	GGT Gly	GCT Ala	TTA Leu	ACT Thr 620	ATT Ile	GAT Asp	r G P A	CA la	ACC Thr	1872			
AAA Lys 625	GAG Glu	ACC Thr	GAG Glu	CAA Gln	GGT Gly 630	Ser	TAT Tyr	ACC Thr	GTA Val	AAT Asn 635	CGT Arg	TTC Phe	GTI Va	A G	BAA Blu	ACC Thr 640	1920			
GGT Gly	AAA Lys	GCA Ala	CTA Leu	CAC His	Glu	GTG Val	ACT Thr	TCA Ser	ACC Thr 650	CAT His	ACC Thr	GCA Ala	. TT	٠ ،	GTG Val	GGC Gly	1968		-	
AAC Asn	CGT Arg	GAA Glu	GAA Glu	Lys	ATA Ile	GAA Glu	TAT Tyr	CGT Arg 665	HIS	AGC Ser	AAT Asn	AAC Asn	CA Gl 67		CAC His	CAT His	2016	٠		
GCC Ala	GGT Gly	TAT Tyr 675	туз	C ACC	AAA Lys	A GAT	ACC Thi	: Бес	AAA Lys	GCT Ala	GTT Val	GAZ Glu 685		A i	ATT Ile	ATC Ile	2064			
Gly	7 Thr 690	s Ser	: His	s Ası	n Asp	69	5 PN	з гъ	, GIY	561	700	)			<u>-</u>	GCC Ala	2112			
Phe 705	e Ası 5	n Gly	y Gl	y As	p G1; 71	y va O	I AS	р тп.	. 116	715	5	, Fis.			<b>-</b> 1	AAT Asn 720	2160			
Ası	p Ar	g Le	u Ph	e Gl 72	y G1 5	у гу	S GI	y AS	730	5	, <u>"</u>	2 7.0	<b>P</b> 0.	- 1	735		2208			
G1	y As	p As	p Ph 74	e Il 0	e As	b GT	y GI	74	5	y ASI	ı AS	<b>P D C</b>	7	50		G GGT G Gly	2256			
GG G1	C AA y Ly	G GG s Gl 75	y As	AT GA	T AT	T TI e Ph	C GI ne Va 76	rr ur	C CG s Ar	T AAI g Ly:	A GG s Gl	C GA y As 76		GT ly	AA! Ası	r GAT n Asp	2304			

ATT Ile	ATT Ile 770	ACC Thr	GAT Asp	TCT Ser	GAC Asp	GGC Gly 775	AAT Asn	GAT Asp	AAA Lys	TTA Leu	TCA Ser 780	TTC Phe	TCT Ser	GAT Asp	TCG Ser	2352
AAC Asn 785	TTA Leu	AAA Lys	GAT Asp	TTA Leu	ACA Thr 790	TTT Phe	GAA Glu	AAA Lys	GTT Val	AAA Lys 795	CAT His	AAT Asn	CTT Leu	GTC Val	ATC Ile 800	2400
ACG Thr	AAT Asn	AGC Ser	AAA Lys	AAA Lys 805	GAG Glu	AAA Lys	GTG Val	ACC Thr	ATT Ile 810	CAA Gln	AAC Asn	TGG Trp	TTC Phe	CGA Arg 815	GAG Glu	2448
GCT Ala	GAT Asp	TTT Phe	GCT Ala 820	AAA Lys	GAA Glu	GTG Val	CCT Pro	AAT Asn 825	TAT Tyr	AAA Lys	GCA Ala	ACT Thr	AAA Lys 830	GAT Asp	GAG Glu	2496
AAA Lys	ATC Ile	GAA Glu 835	Glu	ATC Ile	ATC Ile	GGT Gly	CAA Gln 840	AAT Asn	GGC Gly	GAG Glu	CGG Arg	ATC Ile 845	ACC Thr	TCA Ser	AAG Lys	2544
CAA Gln	GTT Val 850	Asp	GAT Asp	CTT Leu	ATC Ile	GCA Ala 855	AAA Lys	GGT Gly	AAC Asn	GGC Gly	AAA Lys 860	ATT Ile	ACC Thr	CAA Gln	GAT Asp	2592
GAG Glu 865	CTA Leu	TCA Ser	AAA Lys	GTT Val	GTT Val 870	Asp	AAC Asn	тат туг	GAA Glu	TTG Leu 875	meu	AAA Lys	CAT His	AGC Ser	AAA Lys 880	2640
AAT Asn	GTG Val	ACA Thi	A AAC Asn	AGC Ser 885	Leu	GAT Asp	AAG Lys	TTA Leu	ATC Ile 890	Ser	TCT Ser	GTA Val	AGT Ser	GCA Ala 895	TTT Phe	2688
ACC Thr	TCC Ser	TC:	FAA 1 RASI 100	1 Asp	TCG Ser	AGA Arg	AAT AST	GTA Val 905	. Let	GTC Val	GCT Ala	CCA Pro	ACT Thr 910		ATG Met	2736
TT( Lev	GA?	CA Gl: 91	n Sei	TTI Lev	A TCT	r TCT	CTI Lev 920	1 GII	A TTT	r GCT e Ala	r AGO a Arg	G GGA G Gly 925		AGC Ser	TCT Ser	2784
TC'	GCC Ala	a Gl	C TG y Cy	C AAA	A AAG S ASI	C TTO n Pho 931	e Pne	TGG Tr	G AAI p Ly:	A ACC	r Pho	C ACC e Thi	C AGO	TGC Cys	3	2829
TA	GGA'	TCC														2838

# (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 943 amino acids
    (B) TYPE: amino acid

  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys

Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly 20 25 30

Asn	Gly		eu 35	Gln	Asp	Le	۷ د	al	Lys 40	Ala	Ala	G	lu (	Glu	Leu 45	Gly	Ile	Glu
Val	Glr 50		rg	Glu	Glu	Ar	g A	sn 55	Asn	Ile	Ala	Т	hr	Ala 60	Gln	Thr	Ser	Leu
Gly 65	Thr	·I	le	Gln	Thr	Al . 7	a I 0	le	Gly	Leu	Thr	G	1u 75	Arg	Gly	Ile	Val	Leu 80
Ser	Ala	a P	ro	Gln	Ile 85	As	рI	Lys	Leu	Leu	Glr 90	ı L	ys	Thr	Lys	Ala	Gly 95	Gln
Ala	Let	ı G	ly	Ser 100	Ala	Gl	u s	Ser	Ile	Val 105	Glı	n P	sn	Ala	Asn	Lys 110	Ala	Lys
Thr	Va		Jeu L15	Ser	Gly	, Il	е (	Gln	Ser 120	Ile	Lei	u G	ly	Ser	Val 125	Leu	Ala	Gly
Met	As 13		Leu	Asp	Gl	ı Al	.a :	Leu 135	Gln	Asn	As	n S	Ser	Asn 140	Gln	His	Ala	Leu
Ala		s A	Ala	Gly	Lei	ı G]	u :	Leu	Thr	Asn	. Se	r l	Leu 155	Ile	Glu	Asn	Ile	Ala 160
Asn	s Se	r'	Val	Lys	Th	r Le 5	eu	Asp	Glu	Phe	Gl 17	0 V	Glu	Gln	Ile	Ser	Gln 175	Phe
Gly	/ Se	er :	Lys	Le:		n A	sn	Ile	Lys	Gl <sub>y</sub> 185	r Le	u	Gly	Thr	Leu	Gly 190	Asp	Lys
Let	ı Ly		Asn 195		e Gl	уG	ly	Leu	Asp 200	Lys	a Al	.a	Gly	Leu	Gly 205	Leu	. Asp	Val
Ile		er 10	Gly	. Le	u Le	u S	er	Gly 215	Ala	Th	c Al	.a	Ala	Leu 220	Val	Leu	Ala	Asp
Ly 22		sn	Ala	a Se	r Th	ır A	1a 30	Lys	Lys	s Vai	1 G	lу	Ala 235	Gly	/ Phe	e Glu	i Lei	Ala 240
As	n G	ln	Va:	l Va	1 GI 24	.y <i>P</i> 15	sn	Ile	e Thi	r Ly	s A.	la 50	Val	. Sei	Se	г Туг	25!	e Leu
Al	a G	ln	Ar	g Va 26	1 A	la A	la	Gl	y Le	u Se 26	r S 5	er	Thi	Gly	y Pro	o Val 270	L Ala	a Ala
Le	u I	le	Al 27	a Se 5	r T	nr V	/al	Se	r Le	u Al O	a I	le	Sea	r Pro	28	u Ala 5	a Ph	e Ala
Gl		le 90	Al	a As	p L	ys 1	?he	As:	n Hi 5	s Al	a L	ys	Se	r Le	u Gl O	u Se	r Ty	r Ala
G] 3(		ırg	Ph	e Ly	ys L	ys :	ւeu 310	Gl	у Ту	r As	sp G	ly	Asj 31	p As 5	n Le	u Le	u Al	a Glu 320
		ln	. Ar	g G	ly T	hr 25	Gly	, Th	r Il	e As	sp A 3	1a 30	Se	r Va	1 Th	r Al	a Il 33	e Asn 5
T	hr 1	Ala	Le	eu A	la A 40	la	Il€	a Al	a Gl	.у G: 3	ly \ 45	/al	Se	r Al	a Al	a Al 35	a Al	a Gly

Ser Val Ile Ala Ser Pro Ile Ala Leu Leu Val Ser Gly Ile Thr Gly 355

Val Ile Ser Thr Ile Leu Gln Tyr Ser Lys Gln Ala Met Phe Glu His 375 Val Ala Asn Lys Ile His Asn Lys Ile Val Glu Trp Glu Lys Asn Asn His Gly Lys Asn Tyr Phe Glu Asn Gly Tyr Asp Ala Arg Tyr Leu Ala Asn Leu Gln Asp Asn Met Lys Phe Leu Leu Asn Leu Asn Lys Glu Leu Gln Ala Glu Arg Val Ile Ala Ile Thr Gln Gln Gln Trp Asp Asn Asn Ile Gly Asp Leu Ala Gly Ile Ser Arg Leu Gly Glu Lys Val Leu Ser 455 Gly Lys Ala Tyr Val Asp Ala Phe Glu Glu Gly Lys His Ile Lys Ala Asp Lys Leu Val Gln Leu Asp Ser Ala Asn Gly Ile Ile Asp Val Ser Asn Ser Gly Lys Ala Lys Thr Gln His Ile Leu Phe Arg Thr Pro Leu 505 Leu Thr Pro Gly Thr Glu His Arg Glu Arg Val Gln Thr Gly Lys Tyr Glu Tyr Ile Thr Lys Leu Asn Ile Asn Arg Val Asp Ser Trp Lys Ile 535 Thr Asp Gly Ala Ala Ser Ser Thr Phe Asp Leu Thr Asn Val Val Gln Arg Ile Gly Ile Glu Leu Asp Asn Ala Gly Asn Val Thr Lys Glu Thr Lys Ile Ile Ala Lys Leu Gly Glu Gly Asp Asp Asn Val Phe Val Gly Ser Gly Thr Thr Glu Ile Asp Gly Gly Glu Gly Tyr Asp Arg Val His Tyr Ser Arg Gly Asn Tyr Gly Ala Leu Thr Ile Asp Ala Thr 615 Lys Glu Thr Glu Gln Gly Ser Tyr Thr Val Asn Arg Phe Val Glu Thr

Gly Lys Ala Leu His Glu Val Thr Ser Thr His Thr Ala Leu Val Gly 655

Asn Arg Glu Glu Lys Ile Glu Tyr Arg His Ser Asn Asn Gln His His 660

Ala Gly Tyr Tyr Thr Lys Asp Thr Leu Lys Ala Val Glu Glu Ile Ile 675 680 685

Gly Thr Ser His Asn Asp Ile Phe Lys Gly Ser Lys Phe Asn Asp Ala 690 695 700 Phe Asn Gly Gly Asp Gly Val Asp Thr Ile Asp Gly Asn Asp Gly Asn 705 710 715 720

Asp Arg Leu Phe Gly Gly Lys Gly Asp Asp Ile Leu Asp Gly Gly Asn 725 730 735

Gly Asp Asp Phe Ile Asp Gly Gly Lys Gly Asn Asp Leu Leu His Gly 740 745 750

Gly Lys Gly Asp Asp Ile Phe Val His Arg Lys Gly Asp Gly Asn Asp 765

Ile Ile Thr Asp Ser Asp Gly Asn Asp Lys Leu Ser Phe Ser Asp Ser

Asn Leu Lys Asp Leu Thr Phe Glu Lys Val Lys His Asn Leu Val Ile 785 790 795 800

Thr Asn Ser Lys Lys Glu Lys Val Thr Ile Gln Asn Trp Phe Arg Glu 805 810 815

Ala Asp Phe Ala Lys Glu Val Pro Asn Tyr Lys Ala Thr Lys Asp Glu 820 825 830

Lys Ile Glu Glu Ile Ile Gly Gln Asn Gly Glu Arg Ile Thr Ser Lys 835 840 845

Gln Val Asp Asp Leu Ile Ala Lys Gly Asn Gly Lys Ile Thr Gln Asp 850 860

Glu Leu Ser Lys Val Val Asp Asn Tyr Glu Leu Leu Lys His Ser Lys 865 870 875 880

Asn Val Thr Asn Ser Leu Asp Lys Leu Ile Ser Ser Val Ser Ala Phe 885 890 895

Thr Ser Ser Asn Asp Ser Arg Asn Val Leu Val Ala Pro Thr Ser Met 900 905 910

Leu Asp Gln Ser Leu Ser Ser Leu Gln Phe Ala Arg Gly Ser Ser Ser

Ser Ala Gly Cys Lys Asn Phe Phe Trp Lys Thr Phe Thr Ser Cys 930 935 940

### (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2817 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 1..2808

# (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

	,	_	-										•			
ATG Met 1	GCT Ala	ACT Thr	GTT Val	ATA Ile 5	GAT Asp	CTA . Leu	AGC Ser	TTC Phe	CCA Pro 10	AAA Lys	ACT Thr	GGG Gly	GCA Ala	AAA Lys 15	AAA Lys	48
ATT Ile	ATC Ile	CTC Leu	TAT Tyr 20	ATT Ile	CCC Pro	CAA Gln	AAT Asn	TAC Tyr 25	CAA Gln	TAT Tyr	GAT Asp	ACT Thr	GAA Glu 30	CAA Gln	GGT Gly	96
AAT Asn	GGT Gly	TTA Leu 35	CAG Gln	GAT Asp	TTA Leu	GTC Val	AAA Lys 40	GCG Ala	GCC Ala	GAA Glu	GAG Glu	TTG Leu 45	GGG Gly	ATT Ile	GAG Glu	144
GTA Val	CAA Gln 50	AGA Arg	GAA Glu	GAA Glu	CGC Arg	AAT Asn 55	AAT Asn	ATT Ile	GCA Ala	ACA Thr	GCT Ala 60	CAA Gln	ACC Thr	AGT Ser	TTA Leu	192
GGC Gly 65	ACG Thr	ATT Ile	CAA Gln	ACC Thr	GCT Ala 70	ATT Ile	GGC Gly	TTA Leu	ACT Thr	GAG Glu 75	CGT Arg	GGC Gly	ATT Ile	GTG Val	TTA Leu 80	240
TCC Ser	GCT Ala	CCA Pro	CAA Gln	ATT Ile 85	GAT Asp	AAA Lys	TTG Leu	CTA Leu	CAG Gln 90	AAA Lys	ACT Thr	AAA Lys	GCA Ala	GGC Gly 95	CAA Gln	288
GCA Ala	TTA Leu	GGT Gly	TCT Ser	Ala	GAA Glu	AGC Ser	ATT Ile	GTA Val 105	GIII	AAT Asn	GCA Ala	AAT Asn	AAA Lys 110	GCC Ala	AAA Lys	336
ACT Thr	GTA Val	TTA Leu	Ser	GGC Gly	ATT Ile	CAA Gln	TCT Ser 120	TIE	TTA Leu	GGC Gly	TCA Ser	GTA Val 125	пси	GCT Ala	GGA Gly	384
ATG Met	GAT Asp	Lev	A GAT 1 Asp	GAG Glu	GCC Ala	TTA Leu 135	GID	AAT Asn	AAC Asn	AGC Ser	AAC Asr	i Gii	CAT His	GCT Ala	CTT	432
GCT Ala	Lys	GC:	r GG( a Gly	TTG Leu	GAG Glu 150	Leu	ACA Thr	AAT Asr	TCA 1 Ser	TTA Leu 155	. 110	GAA Glu	AA A a Asr	ATI 1 Ile	GCT Ala 160	480
		A GT	a AAI l Ly:	A ACA	Leu	GAC Asp	GAF Glu	A TTT	r GG7 e Gly	GIL	CA Gli	A ATT	r AG1 e Sei	CAA Glr 175	TTT Phe	528
GG' G1	T TC	A AA r Ly	A CT s Le	u Gli	A AAT n Asr	T ATO	AA Lys	A GGG	y ne	A GGC	AC'	r TTI r Le	A GGI 1 Gly 19	1	Lys	576
CT <sup>c</sup>	C AA u Ly	A AA s As 19	n Il	C GG e Gl	r GGA y Gl	A CT	r GA' 1 As 20	Б гА	A GC s Al	T GG( a Gl	C CT y Le	T GG' u Gl; 20	y 20	A GA' u Asj	r GTT o Val	624
AT Il	C TC e Se 21	r Gl	G CT y Le	A TT.	A TC	G GGG r Gl	<b>À У</b> Т	A AC a Th	A GC r Al	T GC	A CT a Le 22	u vu	A CT 1 Le	T GC	A GAT a Asp	672
AA Ly 22	s As	T GC n Al	T TC .a Se	A AC	A GC r Al 23	а ьу	A AA s Ly	A GT s Va	G GG	T GC y Al 23	a Gi	T TT y Ph	T GA e Gl	A TT u Le	G GCA u Ala 240	•

•••							1	1										
•	•			•	1	•						. ,					•	
	AAC Asn	CAA Gln	GTT Val	GTT Val	GGT Gly 245	AAT Asn	ATT Ile	ACC Thr	Lys .	GCC Ala 250	GTT Val	TCT Ser	TCT Ser	TAC Tyr	ATT Ile 255	nea	768	
	GCC Ala	CAA Gln	CGT Arg	GTT Val 260	GCA Ala	GCA Ala	GGT Gly	Leu	TCT Ser 265	TCA Ser	ACT Thr	GGG Gly	CCT Pro	GTG Val 270	AIA	GCT Ala	816	
	TTA Leu	ATT Ile	GCT Ala 275	Ser	ACT Thr	GTT Val	TCT Ser	CTT Leu 280	GCG Ala	ATT	AGC Ser	CCA Pro	TTA Leu 285	ALA	TTT Phe	GCC Ala	864	
	GGT · Gly	ATT Ile 290	Ala	GAT Asp	AAA Lys	TTT	AAT Asn 295	CAT His	GCA Ala	AAA Lys	AGT Ser	TTA Leu 300	Giu	AGT Ser	TAT	GCC Ala	912	
	Glu 305	Arg	, Phe	: Гуз	Lys	TTA Leu 310	Gly	Tyr	Asp	GIY	315	, ASII	L Deu	пец	. Ala	320	960	
	Tyr	Gln	n Arg	g Gly	7 Thr 325	GIY	Tnr	ire	a Asp	330	Jer	. vai		, ,,,,,,,	335	•	1008	
	Thr	: Ala	a Leu	ı Ala 340	a Ala O	a Ile	e Ala	GIY	345	val	. Ser	Ala	a Alu	350	0	GGC Gly	. 1056	
	Ser	c Val	1 Ile 355	e Ala 5	a Ser	r Pro	) lle	360	)	i pen	ı vaı		365	5		C GGT r Gly	1104	
	Val	1 Ile 370	e Ser 0	r Thi	r Il∈	e Leu	375	i Tyr	r ser	г пуз	5 GIII	380	0			G CAC u His	1152	
	Val 385	1 Ala 5	a Ası	n Lys	s Ile	390	S ASI	и пуз	5 110	e var	395	5		2		T AAT n Asn 400	1200	
	His	s Gl	у Гу	s Asi	n Ty:	r Phe	e GIV	u ASI	n Gry	410	0	p mr		J -1	41		1248	
	Ası	n Le	eu Gl	n As 42	p As: 20	n Me	c ràs	S PII6	425	5	u Asi	20	, u 110.	43	30	G TTA		
	Gl	n Al	la Gl 43	lu Ar 35	rg Va	al II	e Ala	.a 110	10	ir Gr	.ii Gi.	01	44	15		AC AAC		
	Il	le Gl 45	ly As 50	sp Le	eu Al	la GI	Ly 11 45	.e se.	st Mr	g ne	,u	46	60			rt AGT eu Ser		
	Gl	GT A# ly Ly 65	AA GC ys Al	C TA	AT GI yr Va	TG GA al As 47	sb ar	G TT	T GA Le Gl	A GA	Lu Oi	GC AF ly Ly 75	A CA	C AT	TT AA le Ly	AA GCC ys Ala 480	1440	
			,		•					-	60 -	-					·	

. '				,								,			:		
											•						
GAT Asp	AAA Lys	, TT ; L€	ra G eu V	Val (	CAG Gln 485	TTG Leu	GAT ' Asp	TCG ( Ser )	Ala .	AAC ( Asn ( 490	GIY	ATT Ile	ATT Ile	GAT Asp	GTG Val 495	AGT Ser	1488
AAT Asn	TCG Ser	: G0	ly I	AAA ( Lys . 500	GCG Ala	AAA Lys	ACT Thr	GIn :	CAT His 505	ATC Ile	TTA Leu	TTC Phe	AGA Arg	ACG Thr 510	CCA Pro	TTA Leu	1536
TTG Leu	ACC Thi	r Pi	CG ( ro (	3GA Gly	ACA Thr	GAG Glu	His	CGT Arg 520	GAA Glu	CGC Arg	GTA Val	CAA Gln	ACA Thr 525	GGT Gly	AAA Lys	TAT Tyr	1584
Glu	Туз 530	r II O	le '	Thr	Lys	CTC Leu	Asn 535	11e	Asn	Arg	vai	540	per	пр	Буз	110	1632
Thr 545	Ası	p G	ly i	Ala	Ala	AGT Ser 550	Ser	Tnr	Pne	Asp	555	1111	ASII	vai	var	560	1680
Arg	Il	e G	3ly	Ile	Glu 565		Asp	ASI	AIA	570	ASII	vai		בינב	575	-3-	1728
Glu	Th	r L	Lys	Ile 580	lle	GCC Ala	гуs 	Leu	585	GIU	GIY	ASP	ASP	590			1776
Val	. Gl	ly S 5	Ser 595	Gly	, Thr	Thr	Glu	600	ASP	GIY	Gry	O.L.	605	-1-			1824
Val	1 Hi 61	is T 10	Tyr	Ser	c Arg	3 GIY	615	Tyr	GIY	Ala	Бец	620				ACC Thr	1872
Lys 625	s G] 5	lu 7	Thr	Glu	u Glm	630	y ser	Tyr	1111	, var	635			, , , , , ,		ACC Thr 640	1920
Gly	у Г	ys i	Ala	Leu	u His	s Glu 5	ı vaı	LIME	. Ser	650	)				65		1968
Ası	n A	rg (	Glu	1 Glu 660	u Lys 0	s Ile	e GIV	ı Tyr	665	5	5 562			67	0	C CAT s His	2016
Al	a G	ly	Tyr 675	r Tyi 5	r Th	ır Ly	s As	680	0	u by.	3 711		68	5		T ATC e Ile	2064
Gl	Lу Т 6	Thr 590	Ser	r Hi	s As	n As	69	5 Pn	е гу	5 GI.	y oc.	70	0			T GCC p Ala	2112
TT Ph 70	ne A	AC Asn	GGT Gl;	y Gl	T GA y As	AT GG sp Gl 71	.y va	C GA	T AC p Th	T AT	T GA e As 71	P 0.	T AA y As	C GA n As	C GG p Gl	C AAT y Asn 720	2160

•

			•		•				1	• • •	. '			1			
GAC (	CGC Arg	TTA Leu	TTT Phe	GGT Gly 725	GGT Gly	AAA ( Lys (	GGC (	Asp	GAT Asp 730	ATT Ile	CTC Leu	GAT Asp	GGT Gly	GGA Gly 735	AAT Asn	2208	
GGT (	GAT Asp	GAT Asp	TTT Phe 740	ATC Ile	GAT Asp	GGC Gly	Gly	AAA Lys 745	GGC Gly	AAC Asn	GAC Asp	CTA Leu	TTA Leu 750	CAC His	GGT Gly	2256	
GGC Gly	AAG Lys	GGC Gly 755	GAT Asp	GAT Asp	ATT Ile	Phe	GTT Val 760	CAC His	CGT Arg	AAA Lys	GGC Gly	GAT Asp 765	GGT Gly	AAT Asn	GAT Asp	2304	
ATT Ile	ATT Ile 770	ACC Thr	GAT Asp	TCT Ser	GAC Asp	GGC Gly 775	AAT Asn	GAT Asp	AAA Lys	TTA Leu	TCA Ser 780	TTC Phe	TCT Ser	GAT Asp	TCG Ser	2352	<i>x</i> -
AAC Asn 785	TTA Leu	AAA Lys	GAT Asp	TTA Leu	ACA Thr 790	TTT Phe	GAA Glu	AAA Lys	GTT Val	AAA Lys 795	CAT His	AAT Asn	CTT Leu	GTC Val	ATC Ile 800	2400	
ACG Thr	AAT Asn	AGC Ser	AAA Lys	AAA Lys 805	GAG Glu	AAA Lys	GTG Val	ACC Thr	ATT Ile 810	CAA Gln	AAC Asn	TGG Trp	TTC Phe	CGA Arg 815	GAG Glu	2448	:
GCT Ala	GAT Asp	TTT Phe	GCT Ala 820	AAA Lys	GAA Glu	GTG Val	CCT Pro	AAT Asn 825	TAT Tyr	AAA Lys	GCA Ala	ACT Thr	AAA Lys 830	GAT Asp	GAG Glu	2496	
AAA Lys	ATC Ile	GAA Glu 835	Glu	ATC Ile	ATC Ile	GGT Gly	CAA Gln 840	AAT Asn	GGC Gly	GAG Glu	CGG Arg	ATC Ile 845	ACC Thr	TCA Ser	AAG Lys	2544	
CAA Gln	GTT Val 850	Asp	GAT Asp	CTT Leu	ATC Ile	GCA Ala 855	AAA Lys	GGT Gly	AAC Asn	GGC Gly	AAA Lys	110	ACC Thr	CAA Gln	GAT Asp	2592	
GAG Glu 865	Leu	TCA Ser	AAA Lys	GTT Val	GTT Val 870	Asp	AAC Asn	TAT Tyr	GAA Glu	TTG Leu 875	шс	AAA Lys	CAT His	AGC Ser	AAA Lys 880	2640	
AAT Asn	GTC Val	ACA L Thi	A AAC Ası	AGC Ser 885	: Leu	GAT Asp	AAG Lys	TTA Leu	A ATO 1 Ile 890	3 261	TCT Ser	GTA Val	AGT Ser	GCA Ala 895	TTT Phe	2688	
ACC Thr	TC(	TC:	AA 7 ASI 900	n Asp	TCC Ser	AGA Arg	AAT Asn	GTA Val 909	r ne	A GTO	GCT L Ala	CCA Pro	ACT Thr 910		A ATG Met	2736	
TTC Lev	GA' LAS	r CA p Gl: 91	n Se	T TTA	A TCT	r TCT r Sei	CT1 Let 920	I GII	A TT	T GCT e Ala	r AGG a Arg	G GGF G Gly 925		CAC Gli	G CAT	2784	
TG( Tr <u>j</u>	AG Se 93	r Ty	C GG r Gl	C CTO y Le	g CG0 u Arg	C CC g Pro 93!	O GIZ	TAI	AGGA	TCC						2817	

### (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 936 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
- Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys
- Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly
- Asn Gly Leu Gln Asp Leu Val Lys Ala Ala Glu Glu Leu Gly Ile Glu
  35 40 45
- Val Gln Arg Glu Glu Arg Asn Asn Ile Ala Thr Ala Gln Thr Ser Leu
- Gly Thr Ile Gln Thr Ala Ile Gly Leu Thr Glu Arg Gly Ile Val Leu
  70 75 80
- Ser Ala Pro Gln Ile Asp Lys Leu Leu Gln Lys Thr Lys Ala Gly Gln
- Ala Leu Gly Ser Ala Glu Ser Ile Val Gln Asn Ala Asn Lys Ala Lys
- Thr Val Leu Ser Gly Ile Gln Ser Ile Leu Gly Ser Val Leu Ala Gly
- Met Asp Leu Asp Glu Ala Leu Gln Asn Asn Ser Asn Gln His Ala Leu 130 135 140
- Ala Lys Ala Gly Leu Glu Leu Thr Asn Ser Leu Ile Glu Asn Ile Ala
  145 150 155 160
- Asn Ser Val Lys Thr Leu Asp Glu Phe Gly Glu Gln Ile Ser Gln Phe 165 170 175
- Gly Ser Lys Leu Gln Asn Ile Lys Gly Leu Gly Thr Leu Gly Asp Lys 180 185 190
- Leu Lys Asn Ile Gly Gly Leu Asp Lys Ala Gly Leu Gly Leu Asp Val
- Ile Ser Gly Leu Leu Ser Gly Ala Thr Ala Ala Leu Val Leu Ala Asp 210 215 220
- Lys Asn Ala Ser Thr Ala Lys Lys Val Gly Ala Gly Phe Glu Leu Ala 225 230 235 240
- Asn Gln Val Val Gly Asn Ile Thr Lys Ala Val Ser Ser Tyr Ile Leu 245 250 255
- Ala Gln Arg Val Ala Ala Gly Leu Ser Ser Thr Gly Pro Val Ala Ala 260 265 270

Leu Ile Ala Ser Thr Val Ser Leu Ala Ile Ser Pro Leu Ala Phe Ala 280 Gly Ile Ala Asp Lys Phe Asn His Ala Lys Ser Leu Glu Ser Tyr Ala 295 Glu Arg Phe Lys Lys Leu Gly Tyr Asp Gly Asp Asn Leu Leu Ala Glu Tyr Gln Arg Gly Thr Gly Thr Ile Asp Ala Ser Val Thr Ala Ile Asn Thr Ala Leu Ala Ala Ile Ala Gly Gly Val Ser Ala Ala Ala Gly Ser Val Ile Ala Ser Pro Ile Ala Leu Leu Val Ser Gly Ile Thr Gly Val Ile Ser Thr Ile Leu Gln Tyr Ser Lys Gln Ala Met Phe Glu His 375 Val Ala Asn Lys Ile His Asn Lys Ile Val Glu Trp Glu Lys Asn Asn His Gly Lys Asn Tyr Phe Glu Asn Gly Tyr Asp Ala Arg Tyr Leu Ala Asn Leu Gln Asp Asn Met Lys Phe Leu Leu Asn Leu Asn Lys Glu Leu 425 Gln Ala Glu Arg Val Ile Ala Ile Thr Gln Gln Gln Trp Asp Asn Asn 440 Ile Gly Asp Leu Ala Gly Ile Ser Arg Leu Gly Glu Lys Val Leu Ser Gly Lys Ala Tyr Val Asp Ala Phe Glu Glu Gly Lys His Ile Lys Ala Asp Lys Leu Val Gln Leu Asp Ser Ala Asn Gly Ile Ile Asp Val Ser 490 Asn Ser Gly Lys Ala Lys Thr Gln His Ile Leu Phe Arg Thr Pro Leu 505 Leu Thr Pro Gly Thr Glu His Arg Glu Arg Val Gln Thr Gly Lys Tyr Glu Tyr Ile Thr Lys Leu Asn Ile Asn Arg Val Asp Ser Trp Lys Ile Thr Asp Gly Ala Ala Ser Ser Thr Phe Asp Leu Thr Asn Val Val Gln Arg Ile Gly Ile Glu Leu Asp Asn Ala Gly Asn Val Thr Lys 570 Glu Thr Lys Ile Ile Ala Lys Leu Gly Glu Gly Asp Asp Asn Val Phe

Val Gly Ser Gly Thr Thr Glu Ile Asp Gly Gly Glu Gly Tyr Asp Arg

Val His Tyr Ser Arg Gly Asn Tyr Gly Ala Leu Thr Ile Asp Ala Thr Lys Glu Thr Glu Gln Gly Ser Tyr Thr Val Asn Arg Phe Val Glu Thr Gly Lys Ala Leu His Glu Val Thr Ser Thr His Thr Ala Leu Val Gly Asn Arg Glu Glu Lys Ile Glu Tyr Arg His Ser Asn Asn Gln His His 665 Ala Gly Tyr Tyr Thr Lys Asp Thr Leu Lys Ala Val Glu Glu Ile Ile Gly Thr Ser His Asn Asp Ile Phe Lys Gly Ser Lys Phe Asn Asp Ala 695 Phe Asn Gly Gly Asp Gly Val Asp Thr Ile Asp Gly Asn Asp Gly Asn Asp Arg Leu Phe Gly Gly Lys Gly Asp Asp Ile Leu Asp Gly Gly Asn Gly Asp Asp Phe Ile Asp Gly Gly Lys Gly Asn Asp Leu Leu His Gly Gly Lys Gly Asp Asp Ile Phe Val His Arg Lys Gly Asp Gly Asn Asp Ile Ile Thr Asp Ser Asp Gly Asn Asp Lys Leu Ser Phe Ser Asp Ser Asn Leu Lys Asp Leu Thr Phe Glu Lys Val Lys His Asn Leu Val Ile Thr Asn Ser Lys Lys Glu Lys Val Thr Ile Gln Asn Trp Phe Arg Glu Ala Asp Phe Ala Lys Glu Val Pro Asn Tyr Lys Ala Thr Lys Asp Glu 825 Lys Ile Glu Glu Ile Ile Gly Gln Asn Gly Glu Arg Ile Thr Ser Lys Gln Val Asp Asp Leu Ile Ala Lys Gly Asn Gly Lys Ile Thr Gln Asp Glu Leu Ser Lys Val Val Asp Asn Tyr Glu Leu Leu Lys His Ser Lys Asn Val Thr Asn Ser Leu Asp Lys Leu Ile Ser Ser Val Ser Ala Phe Thr Ser Ser Asn Asp Ser Arg Asn Val Leu Val Ala Pro Thr Ser Met 905 Leu Asp Gln Ser Leu Ser Ser Leu Gln Phe Ala Arg Gly Ser Gln His 920 Trp Ser Tyr Gly Leu Arg Pro Gly

#### (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2861 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear (ii) MOLECULE TYPE: DNA (genomic) (ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1..2853 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13: ATG GCT ACT GTT ATA GAT CTA AGC TTC CCA AAA ACT GGG GCA AAA AAA 48 Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys ATT ATC CTC TAT ATT CCC CAA AAT TAC CAA TAT GAT ACT GAA CAA GGT 96 Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly 20 AAT GGT TTA CAG GAT TTA GTC AAA GCG GCC GAA GAG TTG GGG ATT GAG 144 Asn Gly Leu Gln Asp Leu Val Lys Ala Ala Glu Glu Leu Gly Ile Glu 35 GTA CAA AGA GAA GAA CGC AAT AAT ATT GCA ACA GCT CAA ACC AGT TTA 192 Val Gln Arg Glu Glu Arg Asn Asn Ile Ala Thr Ala Gln Thr Ser Leu 55 GGC ACG ATT CAA ACC GCT ATT GGC TTA ACT GAG CGT GGC ATT GTG TTA 240 Gly Thr Ile Gln Thr Ala Ile Gly Leu Thr Glu Arg Gly Ile Val Leu TCC GCT CCA CAA ATT GAT AAA TTG CTA CAG AAA ACT AAA GCA GGC CAA 288 Ser Ala Pro Gln Ile Asp Lys Leu Leu Gln Lys Thr Lys Ala Gly Gln GCA TTA GGT TCT GCC GAA AGC ATT GTA CAA AAT GCA AAT AAA GCC AAA 336 Ala Leu Gly Ser Ala Glu Ser Ile Val Gln Asn Ala Asn Lys Ala Lys 105 ACT GTA TTA TCT GGC ATT CAA TCT ATT TTA GGC TCA GTA TTG GCT GGA 384 Thr Val Leu Ser Gly Ile Gln Ser Ile Leu Gly Ser Val Leu Ala Gly 120 ATG GAT TTA GAT GAG GCC TTA CAG AAT AAC AGC AAC CAA CAT GCT CTT 432 Met Asp Leu Asp Glu Ala Leu Gln Asn Asn Ser Asn Gln His Ala Leu 135 GCT AAA GCT GGC TTG GAG CTA ACA AAT TCA TTA ATT GAA AAT ATT GCT 480 Ala Lys Ala Gly Leu Glu Leu Thr Asn Ser Leu Ile Glu Asn Ile Ala 150 AAT TCA GTA AAA ACA CTT GAC GAA TTT GGT GAG CAA ATT AGT CAA TTT Asn Ser Val Lys Thr Leu Asp Glu Phe Gly Glu Gln Ile Ser Gln Phe 165 170

(2) INFORMATION FOR SEQ ID NO:13:

															•			•
			,	ı	•	•						• •			;			
GG' Gl'	r T y S	CA ler l	Lys	CTA Leu 180	CAA Gln	AAT Asn	ATC Ile	гàг	GGC ' Gly :	TTA ( Leu	GGG . Gly	ACT Thr	neu	GGA Gly 190	GAC Asp	AAĀ Lys	576	
CT( Le	C A u I	lys /		<b>&gt;</b>	GGT Gly	GGA Gly	CTT Leu	GAT Asp 200	AAA Lys	GCT Ala	GGC Gly	CTT Leu	GGT Gly 205	TTA Leu	GAT Asp	GTT Val	624	
Il	e S	Ser ( 210	Glý	Leu	Leu	TCG Ser	G1y 215	Ата	Thr	Ala	Ala	220	Val	Dea	7124		672	
Ly 22	s <i>I</i>	Asn .	Ala	Ser	Thr	GCT Ala 230	ьуs	гÀг	vai	GIY	235	GIY	1110	014		240	720	
AA As	C (	CAA Gln	GTT Val	GTT Val	GGT Gly 245	AAT Asn	ATT Ile	ACC Thr	AAA Lys	GCC Ala 250	GTT Val	TCT Ser	TCT Ser	TAC Tyr	ATT Ile 255	TTA Leu	768	
Al	la (	Gln	Arg	Val 260	Ala	GCA Ala	GIY	Leu	265	261	1111	OL,		270			816	
L€	eu	Ile	Ala 275	Ser	r Thr	r GTT val	. Ser	280	Ala	116	561	110	285				864	
G.	ly	Ile 290	Ala	a Asp	o Lys	A TTT s Phe	295	HIS	ALA	. шуз	Ser	300	014	-	-1-		912	
G:	1u 05	Arg	Phe	e Lys	s,Lys	A TTA s Leu 310	n Gry	, làr	Asp	, Giy	315	, 1101				320	960	
T	yr	Gln	Arg	g Gl	y Thi 32!		y Thi	2 116	: Asp	330	)				335		1008	
Т	hr	Ala	a Leu	u Ala 340	a Ala	a IIe	e Ala	a GI	345	5	. 501			350	)	GGC Gly	1056	
S	Ser	Val	1 Ile 35	e Ala	a Se	er Pro	0 110	360	0	u nec	u va.		36	5	-	GGT Gly	1104	
V	/al	. Ile	e Sė: 0	r Th	r II	le Le	37	n 1y. 5	ı se.	L Lly.	5 012	38	0			G CAC	1152	
7	Val 385	. Ala	a As	n Ly	ys Il	le H1 39	o O	in Ly	S 11	e va	39	5	F	1		T AAT n Asn 400	1200	
]	CAC His	: GGT ; Gl	r AA y Ly	.G AA /s As	sn Ty	AC TT yr Ph 05	T GA ie Gl	A AA .u As	T GG n Gl	T TA y Ty 41	_ A3	T GC p Al	C CG a Ar	T TA g Ty	T CT r Le 41	T GCG u Ala 5	1248	

.

AAT Asn	TTA Leu	CAJ Gl:	n As	sp A	AAT . Asn	ATG Met	AAA Lys	Phe	Leu	CTG Leu	AAC Asn	TTA Leu	AAC Asn	AAA Lys 430	GAG Glu	TTA Leu	1296
CAG Gln	GCA Ala	A GAA A Glu 435	A CG	20 GT G rg V	GTC Val	ATC Ile	Ala	АТТ	ACT Thr	CAG Gln	CAG Gln	CAA Gln	TGG Trp 445	GAT	AAC Asn	AAC Asn	1344
ATT Ile	GGT Gly 450	/ Asp	T TT p Le	TA (	GCT Ala	Gly	ATT Ile 455	AGC Ser	CGT Arg	TTA Leu	GGT Gly	GAA Glu 460	AAA Lys	GTC Val	CTT Leu	AGT Ser	1392
GGT Gly 465	Lys	. GC(	C T/	AT (	Val	GAT Asp 470	GCG Ala	TTT Phe	GAA Glu	GAA Glu	GGC Gly 475	гуѕ	CAC His	ATT Ile	AAA Lys	GCC Ala 480	1440
GAT Asp	AAA Lys	TT.	A G'	/al (	CAG Gln 485	Leu	GAT Asp	TCG Ser	GCA Ala	AAC Asn 490	GGT Gly	ATT Ile	ÄTT Ile	GAT Asp	GTG Val 495	Ser	1488
AAT Asn	TCC Se:	₃ GG r Gl	ly L	AAA Lys 500	GCG Ala	AAA Lys	ACT Thr	CAG Gln	CAT His 505	: TTE	TTA Leu	TTC Phe	AGA Arg	ACG Thr 510	PIO	TTA Leu	1536
TTG Leu	AC(	G CC r Pr 51	ro G	GA 31y	ACA Thr	GAG Glu	CAT His	CGT Arg 520	GIU	, CGC , Arg	GTA Val	CAA Gln	ACA Thr 525	GLY	AAA Lys	TAT	1584
GAA Glu	TA:	r Il	rt A le I	\CC Thr	AAG Lys	CTC Leu	AAT Asn 535	ı ıre	AAC Asn	: CGT	r GTA J Val	A GAT L Asp 540	JOL	TGG Trp	AAA Lys	ATT	1632
ACA Thr 545	c As	T G( p G	ST G ly I	3CA Ala	GCA Ala	AGT Ser 550	s ser	ACC Thr	: TTT : Phe	: GAT	r TTA p Leu 555	<u> </u>	AAC Asn	: GTT ı Val	GTI Val	CAG Gln 560	1680
CG1 Arg	: AT ; Il	T G( e G	GT <i>I</i> ly :	ATT Ile	GAA Glu 565	ı Leu	. GAC ı Ası	: AAT Asr	' GCT 1 Ala	T GGF a Gly 570	y ASI	: GTA 1 Val	ACT Thr	: AAA : Lys	A ACC 5 Thr 575	AAA Lys	1728
GA <i>I</i> Glı	A AC	A A	.ys ]	ATT Ile 580	lle	GCC Ala	: AAF a Lys	L CTI	T GGT u Gly 585	у Ст	A GG! u Gl	[ GA] y Asj	GAC Ası	C AAC p Asr 590		A TTT L Phe	1776
GT: Va	r G0 1 GJ	ly S	TCT ( Ser (	GGT Gly	ACG Thr	; ACC	∃ GAJ r Glı	A ATI u Ile 600	e Ası	r GG p Gl	c GG′ y Gl′	r GAJ y Gl	A GGT u Gly 605	ž - 3,	C GA( r As)	C CGA p, Arg	1824
GT' Va	l Hi	AC T. is T	'AT Tyr	AGC Ser	: CGT	r GGI	A AAG y Ası 61	n Tyi	Г GG' r Gl	T GC y Al	T TT/ a Le	A ACT		r GA' e Asi	T GC	A ACC a Thr	1872
AAI Ly: 62	rs G]	AG A lu I	ICC [hr	GAG Glu	; CAJ 1 Gl:	A GGT n Gly 630	y se:	T TA' r Ty	T AC r Th	C GT	TA AA' al As: 63	in Pr	r TT g Ph	C GT. e Va	A GA 1 Gl	A ACC u Thr 640	
GG Gl	T A	AA G	3CA Ala	CTA Let	A CAC u His	s GI1	A GT u Va	G AC	T TC	CA AC er Th 65	ir ni	T AC	C GC r Al	A TT	'A GT eu Va 65	G GGC 1 Gly 5	1968

. •				-	'n						. '						
										,							
AAC Asn	CGT Arg	GAA Glu	GAA Glu 660	AAA Lys	ATA Ile	GAA ' Glu	Tyr 1	CGT Arg	CAT His	AGC Ser	AAT Asn	AAC Asn	CAG Gln 670	CAC His	CAT His	2016	
GCC Ala	GGT Gly	TAT Tyr 675	TAC Tyr	ACC Thr	AAA Lys	Asp	ACC Thr	rTG Leu	AAA Lys	GCT Ala	GTT Val	GAA Glu 685	GAA Glu	ATT Ile	ATC Ile	2064	
GGT Gly	ACA Thr 690	TCA Ser	CAT His	AAC Asn	GAT Asp	ATC Ile 695	TTT . Phe	AAA Lys	GGT Gly	AGT Ser	AAG Lys 700	TTC Phe	AAT Asn	GAT Asp	GCC Ala	2112	
TTT Phe 705	AAC Asn	GGT Gly	GGT Gly	GAT Asp	GGT Gly 710	GTC Val	GAT Asp	ACT Thr	ATT Ile	GAC Asp 715	GGT Gly	AAC Asn	GAC Asp	GGC Gly	AAT Asn 720	2160	
GAC Asp	CGC Arg	TTA Leu	TTT Phe	GGT Gly 725	GGT Gly	AAA Lys	GGC Gly	GAT Asp	GAT Asp 730	ATT Ile	CTC Leu	GAT Asp	GGT Gly	GGA Gly 735	AAT Asn	2208	
GGT Gly	GAT Asp	GAT Asp	TTT Phe 740	ATC Ile	GAT Asp	GGC Gly	GGT Gly	AAA Lys 745	GGC Gly	AAC Asn	GAC Asp	CTA Leu	TTA Leu 750	CAC His	GGT Gly	2256	
GGC Gly	AAG Lys	GGC Gly 755	Asp	GAT Asp	ATT Ile	TTC Phe	GTT Val 760	CAC His	CGT	AAA Lys	GGC Gly	GAT Asp 765	GGT Gly	AAT Asn	GAT Asp	2304	
ATT Ile	ATT Ile 770	Thr	GAT Asp	TCT Ser	GAC Asp	GGC Gly 775	AAT Asn	GAT Asp	AAA Lys	TTA Leu	TCA Ser 780	TTC Phe	TCT Ser	GAT Asp	TCG Ser	2352	
AAC Asn 785	Leu	AAA Lys	A GAT	TTA Leu	ACA Thr 790	Pne	GAA Glu	AAA Lys	GTT Val	AAA Lys 795	****	AAT Asn	CTT Leu	GTC Val	ATC Ile 800	2400	
ACG Thr	AAT Asr	AGC 1 Sei	C AAA	A AAA E Lys 805	GIU	AAA Lys	GTG Val	ACC Thr	ATT Ile		AAC Asn	TGG Trp	TTC Phe	CGA Arg 815	GAG Glu	2448	
Ala	ASP	Pho	e Ala 820	a Lys O	GIU	ı vaı	PIO	825	. 191				830	)	GAG Glu	2496	•
Lys	s Ile	e Gl <sup>.</sup> 83	u Gli 5	u Ile	e 116	s GTŽ	840	ASI	ı Gı	, 01.		845	5	,	AAG Lys	2544	
Glr	n Va 85	l As O	p As	p Le	u IIe	855	r rys	GI)	y As		86	0			A GAT n Asp	2592	
Gl: 86	u Le 5	u Se	r Ly	s Va	1 Va. 87	O AS	) ASI	1 IY.	L GI	87	5	,			C AAA C Lys 880	2640	
AA' As	T GT n Va	G AC	A AA ir As	C AG n Se 88	r Le	A GA' u As	r AAG p Lys	G TT.	A AT u Il 89	e 50	A TC r Se	T GT. r Va	A AG' l Se	T GC. r Al 89	A TTT a Phe 5	2688	
						•			-	69 -							

ACC TCG TCT AAT GAT TCG AGA AAT GTA TTA GTG GCT CCA ACT TCA ATG 2736 Thr Ser Ser Asn Asp Ser Arg Asn Val Leu Val Ala Pro Thr Ser Met 905 TTG GAT CAA AGT TTA TCT TCT CTT CAA TTT GCT AGG GGA TCT TGC AAC 2784 Leu Asp Gln Ser Leu Ser Ser Leu Gln Phe Ala Arg Gly Ser Cys Asn ATT GTG CCT GTG AGC ATT GTG AGC CGC AAC ATT GTG TAC ACC CGC GCG 2832 Ile Val Pro Val Ser Ile Val Ser Arg Asn Ile Val Tyr Thr Arg Ala 935 CAA CCT AAC CAA GAC ATT GTG TAGGATCC 2861 Gln Pro Asn Gln Asp Ile Val (2) INFORMATION FOR SEQ ID.NO:14: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 951 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14: Met Ala Thr Val Ile Asp Leu Ser Phe Pro Lys Thr Gly Ala Lys Lys Ile Ile Leu Tyr Ile Pro Gln Asn Tyr Gln Tyr Asp Thr Glu Gln Gly Asn Gly Leu Gln Asp Leu Val Lys Ala Glu Glu Leu Gly Ile Glu Val Gln Arg Glu Glu Arg Asn Asn Ile Ala Thr Ala Gln Thr Ser Leu Gly Thr Ile Gln Thr Ala Ile Gly Leu Thr Glu Arg Gly Ile Val Leu Ser Ala Pro Gln Ile Asp Lys Leu Leu Gln Lys Thr Lys Ala Gly Gln Ala Leu Gly Ser Ala Glu Ser Ile Val Gln Asn Ala Asn Lys Ala Lys Thr Val Leu Ser Gly Ile Gln Ser Ile Leu Gly Ser Val Leu Ala Gly 120 Met Asp Leu Asp Glu Ala Leu Gln Asn Asn Ser Asn Gln His Ala Leu Ala Lys Ala Gly Leu Glu Leu Thr Asn Ser Leu Ile Glu Asn Ile Ala 155 Asn Ser Val Lys Thr Leu Asp Glu Phe Gly Glu Gln Ile Ser Gln Phe 170 Gly Ser Lys Leu Gln Asn Ile Lys Gly Leu Gly Thr Leu Gly Asp Lys

Leu Lys Asn Ile Gly Gly Leu Asp Lys Ala Gly Leu Gly Leu Asp Val Ile Ser Gly Leu Leu Ser Gly Ala Thr Ala Ala Leu Val Leu Ala Asp Lys Asn Ala Ser Thr Ala Lys Lys Val Gly Ala Gly Phe Glu Leu Ala Asn Gln Val Val Gly Asn Ile Thr Lys Ala Val Ser Ser Tyr Ile Leu Ala Gln Arg Val Ala Ala Gly Leu Ser Ser Thr Gly Pro Val Ala Ala 265 Leu Ile Ala Ser Thr Val Ser Leu Ala Ile Ser Pro Leu Ala Phe Ala 280 Gly Ile Ala Asp Lys Phe Asn His Ala Lys Ser Leu Glu Ser Tyr Ala 295 Glu Arg Phe Lys Lys Leu Gly Tyr Asp Gly Asp Asn Leu Leu Ala Glu Tyr Gln Arg Gly Thr Gly Thr Ile Asp Ala Ser Val Thr Ala Ile Asn Thr Ala Leu Ala Ala Ile Ala Gly Gly Val Ser Ala Ala Ala Ala Gly Ser Val Ile Ala Ser Pro Ile Ala Leu Leu Val Ser Gly Ile Thr Gly Val Ile Ser Thr Ile Leu Gln Tyr Ser Lys Gln Ala Met Phe Glu His Val Ala Asn Lys Ile His Asn Lys Ile Val Glu Trp Glu Lys Asn Asn His Gly Lys Asn Tyr Phe Glu Asn Gly Tyr Asp Ala Arg Tyr Leu Ala 410 Asn Leu Gln Asp Asn Met Lys Phe Leu Leu Asn Leu Asn Lys Glu Leu Gln Ala Glu Arg Val Ile Ala Ile Thr Gln Gln Gln Trp Asp Asn Asn 440 Ile Gly Asp Leu Ala Gly Ile Ser Arg Leu Gly Glu Lys Val Leu Ser Gly Lys Ala Tyr Val Asp Ala Phe Glu Glu Gly Lys His Ile Lys Ala 470 Asp Lys Leu Val Gln Leu Asp Ser Ala Asn Gly Ile Ile Asp Val Ser 490 Asn Ser Gly Lys Ala Lys Thr Gln His Ile Leu Phe Arg Thr Pro Leu

Leu Thr Pro Gly Thr Glu His Arg Glu Arg Val Gln Thr Gly Lys Tyr

Glu Tyr Ile Thr Lys Leu Asn Ile Asn Arg Val Asp Ser Trp Lys Ile Thr Asp Gly Ala Ala Ser Ser Thr Phe Asp Leu Thr Asn Val Val Gln 555 Arg Ile Gly Ile Glu Leu Asp Asn Ala Gly Asn Val Thr Lys Glu Thr Lys Ile Ile Ala Lys Leu Gly Glu Gly Asp Asp Asn Val Phe 585 Val Gly Ser Gly Thr Thr Glu Ile Asp Gly Gly Glu Gly Tyr Asp Arg Val His Tyr Ser Arg Gly Asn Tyr Gly Ala Leu Thr Ile Asp Ala Thr Lys Glu Thr Glu Gln Gly Ser Tyr Thr Val Asn Arg Phe Val Glu Thr 635 Gly Lys Ala Leu His Glu Val Thr Ser Thr His Thr Ala Leu Val Gly Asn Arg Glu Glu Lys Ile Glu Tyr Arg His Ser Asn Asn Gln His His Ala Gly Tyr Tyr Thr Lys Asp Thr Leu Lys Ala Val Glu Glu Ile Ile Gly Thr Ser His Asn Asp Ile Phe Lys Gly Ser Lys Phe Asn Asp Ala Phe Asn Gly Gly Asp Gly Val Asp Thr Ile Asp Gly Asn Asp Gly Asn Asp Arg Leu Phe Gly Gly Lys Gly Asp Asp Ile Leu Asp Gly Gly Asn Gly Asp Asp Phe Ile Asp Gly Gly Lys Gly Asn Asp Leu Leu His Gly Gly Lys Gly Asp Asp Ile Phe Val His Arg Lys Gly Asp Gly Asn Asp Ile Ile Thr Asp Ser Asp Gly Asn Asp Lys Leu Ser Phe Ser Asp Ser Asn Leu Lys Asp Leu Thr Phe Glu Lys Val Lys His Asn Leu Val Ile Thr Asn Ser Lys Lys Glu Lys Val Thr Ile Gln Asn Trp Phe Arg Glu Ala Asp Phe Ala Lys Glu Val Pro Asn Tyr Lys Ala Thr Lys Asp Glu 825 Lys Ile Glu Glu Ile Ile Gly Gln Asn Gly Glu Arg Ile Thr Ser Lys Gln Val Asp Asp Leu Ile Ala Lys Gly Asn Gly Lys Ile Thr Gln Asp

Glu Leu Ser Lys Val Val Asp Asn Tyr Glu Leu Leu Lys His Ser Lys 870

Asn Val Thr Asn Ser Leu Asp Lys Leu Ile Ser Ser Val Ser Ala Phe 890 885

Thr Ser Ser Asn Asp Ser Arg Asn Val Leu Val Ala Pro Thr Ser Met

Leu Asp Gln Ser Leu Ser Ser Leu Gln Phe Ala Arg Gly Ser Cys Asn 925 920

Ile Val Pro Val Ser Ile Val Ser Arg Asn Ile Val Tyr Thr Arg Ala 935

Gln Pro Asn Gln Asp Ile Val 945

- (2) INFORMATION FOR SEQ ID NO:15:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 6 amino acids
    - (B) TYPE: amino acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: peptide
  - (ix) FEATURE:
    - (A) NAME/KEY: Modified-site
    - (B) LOCATION: 3
    - (D) OTHER INFORMATION: /note= "The amino acid at this location can be either Lys, Asp, Val or Asn."
  - (ix) FEATURE:
    - (A) NAME/KEY: Modified-site
    - (B) LOCATION: 5
    - (D) OTHER INFORMATION: /note= "The amino acid at this location can be either Lys, Asp, Val or Asn."
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gly Gly Xaa Gly Xaa Asp 5